Search for the SM Higgs boson in the di-tau final state at Tevatron

Pierluigi Totaro,

University of Trieste

On behalf of the

CDF and DØ collaborations

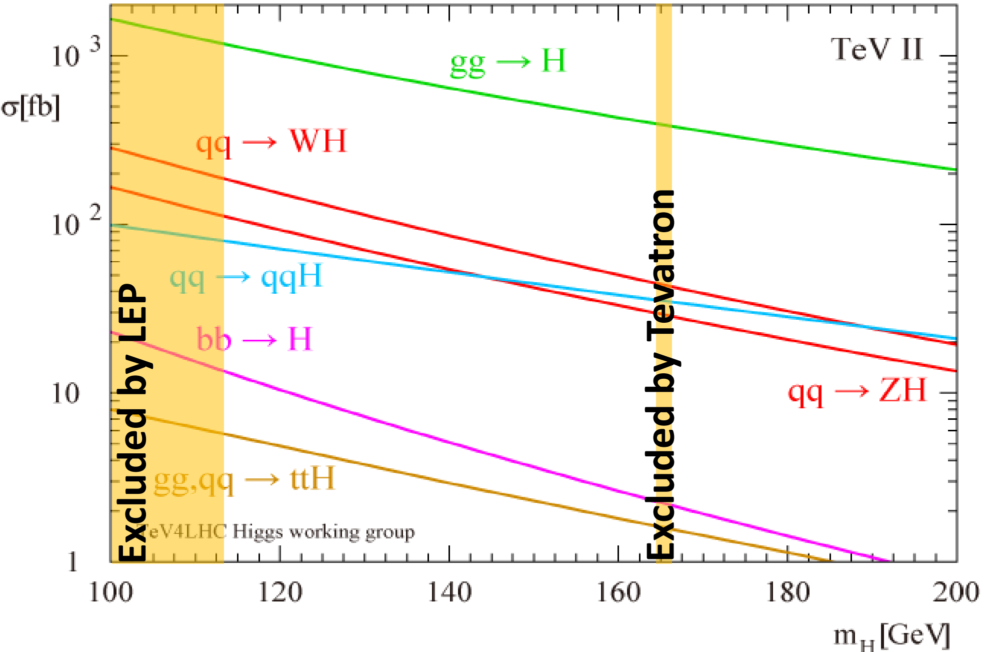
35th International Conference on High Energy Physics

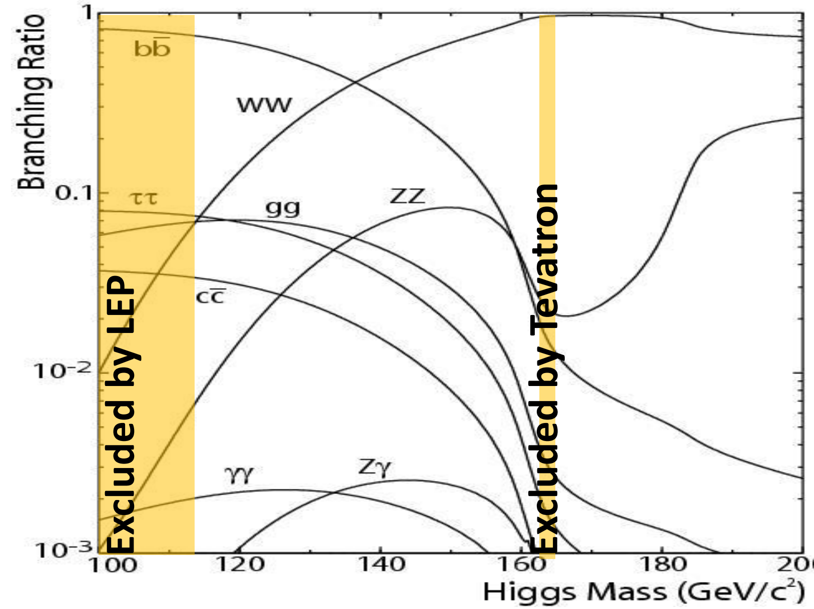
Paris, July 23rd 2010

Outline

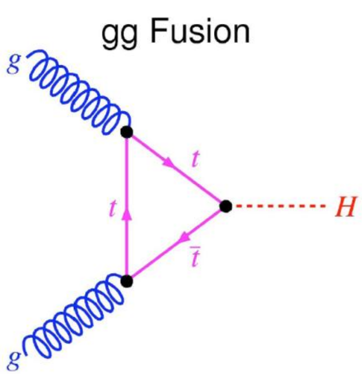
- Standard Model Higgs production and decay at Tevatron
  - Low mass searches
- Motivation of the $H \rightarrow \tau\tau$ searches
- Analysis strategies for CDF and DØ experiments
- Results: CDF 2.3 fb^{-1} DØ 4.9 fb^{-1}
- Conclusions

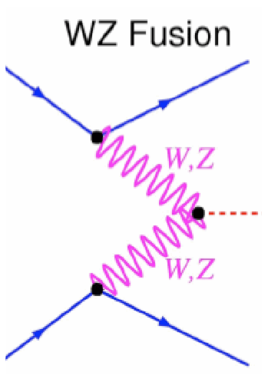
Higgs production and decay at Tevatron

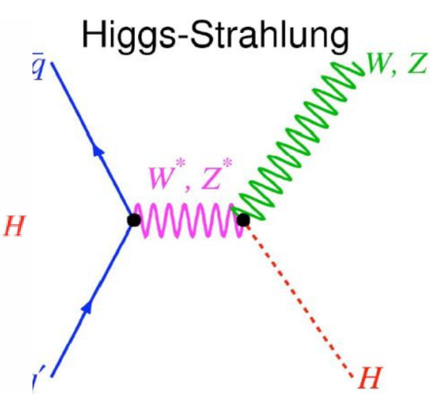




Primary production modes:







1.8 – 0.2 pb   0.1 – 0.02 pb   0.5 – 0.03 pb

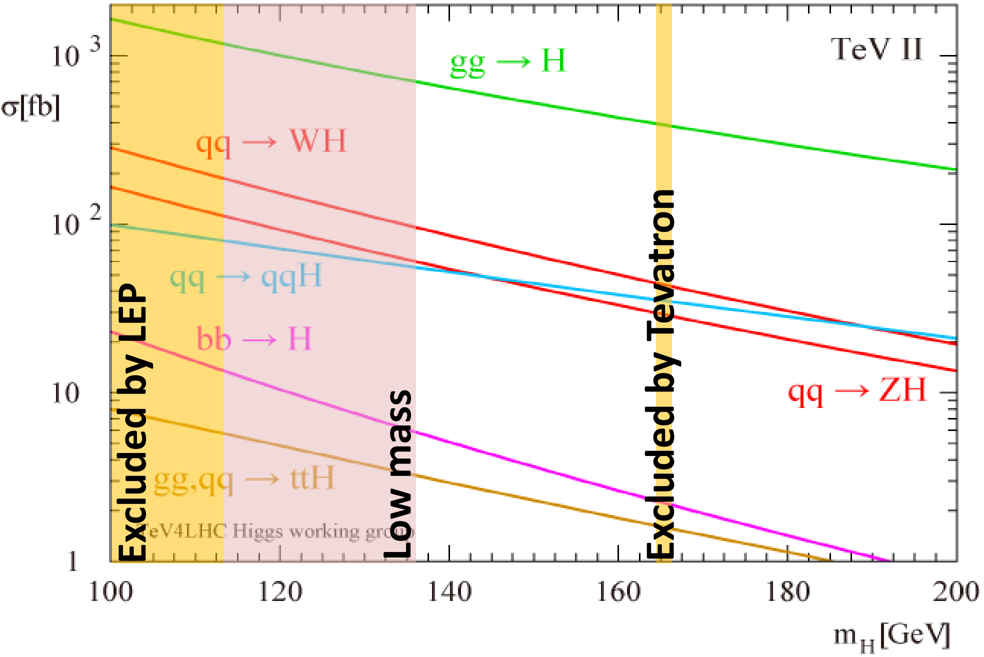
Higgs **direct production dominates** in the whole mass range probed at Tevatron.

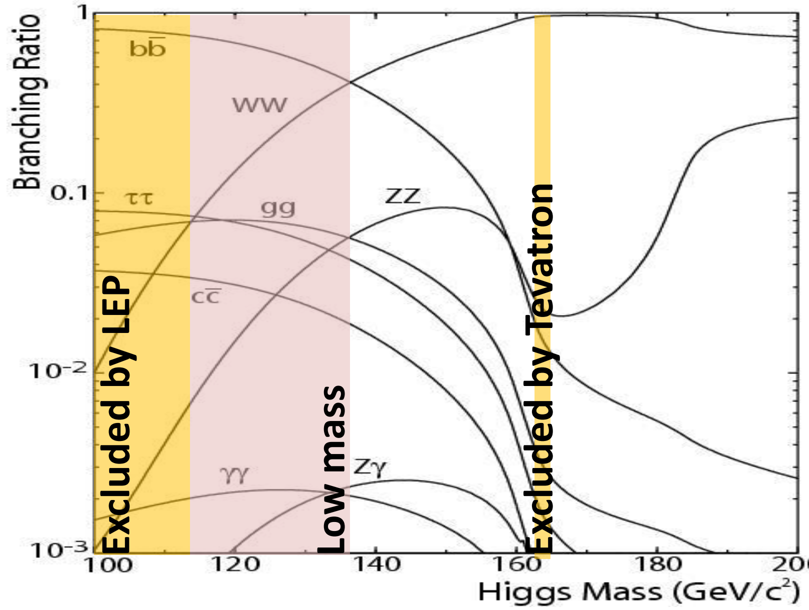
Principal decay modes:

H→bb for M_H<135 GeV/c^2

H → WW* for M_H>135 GeV/c^2

Low Mass Higgs searches at Tevatron





Low mass Higgs (M_H < 135 GeV/c^2)

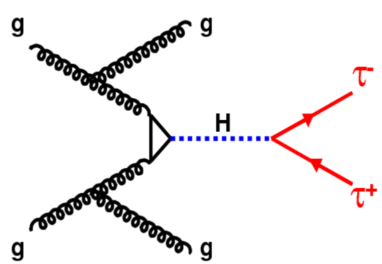
1) gg→H→bb overwhelmed by QCD multijet processes
2) WH→lνbb, ZH→ννbb, ZH→llbb (associated production)
   leptonic decays of W/Z and b-tagging allow to keep bkg under control
3) **May additional secondary channels help in the Higgs hunting?**
   H → ττ is a complementary process with no negligible B.R.

H→ττ searches: motivation

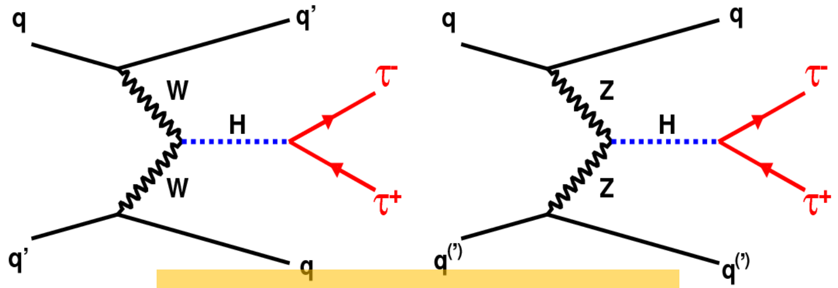
H→ττ branching ratio is small(<10%)

BUT

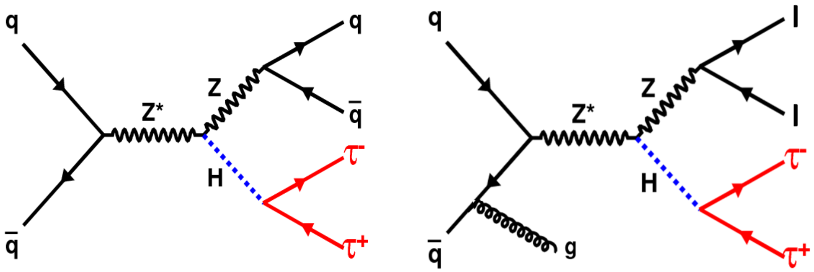
1)Different channels can be studied simultaneously
2)Direct production and VBF become accesible
3)Hadronic W/Z decays in the associated production can be considered



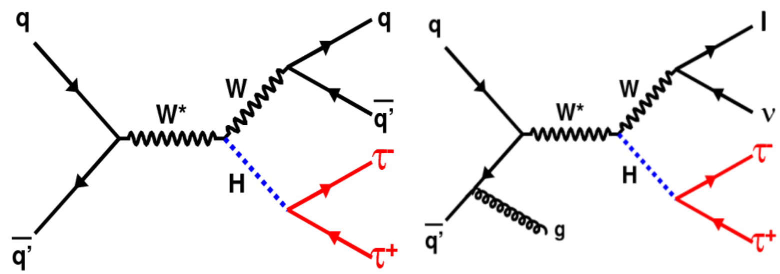
GLUON FUSION



VECTOR BOSON FUSION



Z BOSON ASSOCIATED PRODUCTION



W BOSON ASSOCIATED PRODUCTION

What about tau leptons?

- Heavy particles: 1.78 GeV/c^2
- Short lived: mean lifetime 291 ps ($c\tau$=87 μm) → Detectable only through their decay products
- Decay modes:
  - $\tau \to \nu_\tau \nu_e e$ (B.R.~17%)
  - $\tau \to \nu_\tau \nu_\mu \mu$ (B.R.~17%)
  - $\tau \to \nu_\tau X_h$ (B.R.~65%)

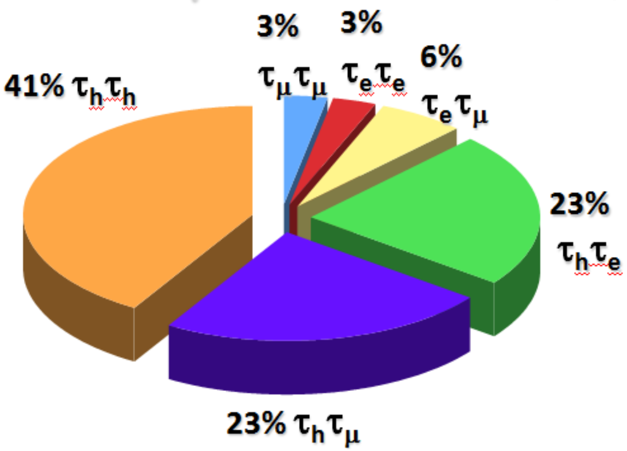
Look for isolated electrons or muons

Hadronic decays:

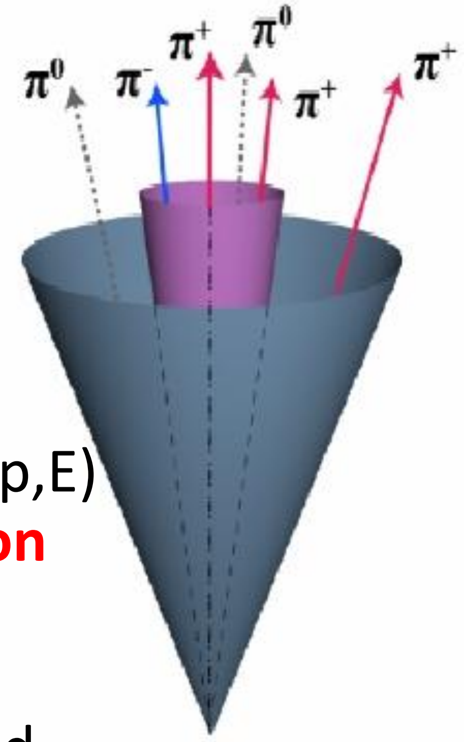
1-prongs $\tau^\pm \to \nu_\tau + h^\pm + n(\pi^0)$

3-prongs $\tau^\pm \to \nu_\tau + h^\pm h^\pm h^\pm + n(\pi^0)$

- Di-tau decay combinations:
  - Hadronic+hadronic: 42 % overwhelming QCD
  - Leptonic+hadronic: 46 % **golden channel**
  - ee/$\mu\mu$: 6 % overwhelming Drell-Yan
  - eμ/μe: 6% clean signature but low B.R.



Hadronic tau identification



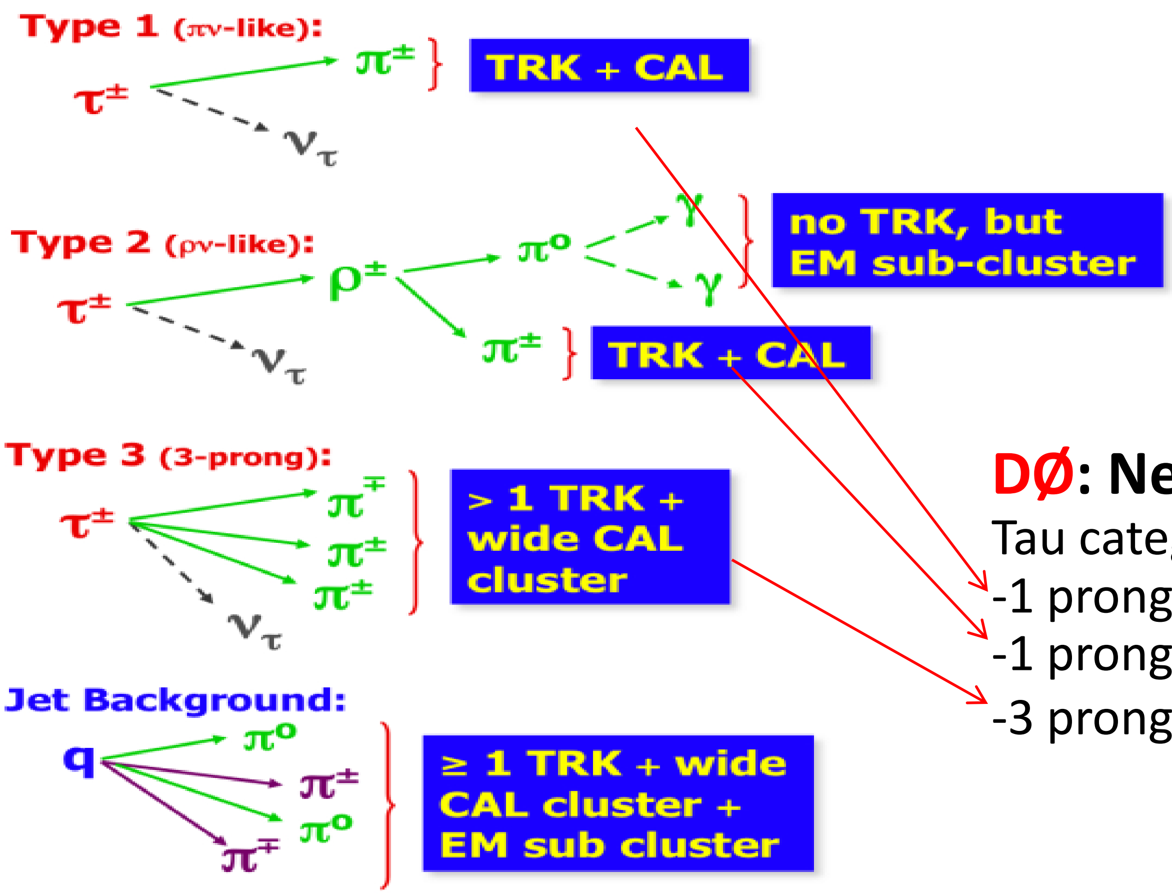
Very challenging task

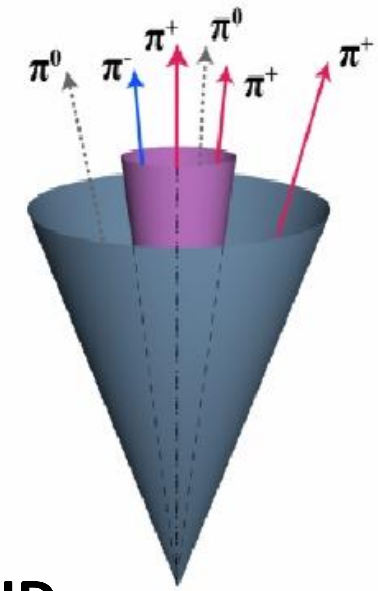
The signature: narrow calorimeter clusters with low multiplicity tracks
QCD jets can easily lead to fakes

Reconstruction: very difficult due to the not detected neutrinos; only the "visible" fraction of the energy can be used to build the $P_{had}(p,E)$

Identification: - based on **calorimeter and track isolation requirements**
- Multivariate selections are better than rectangular cut to exploit correlations and provide a good **τ-jet separation**
- best performances achieved by considering separately **different tau categories**

Hadronic tau identification





DØ: Neural network ID

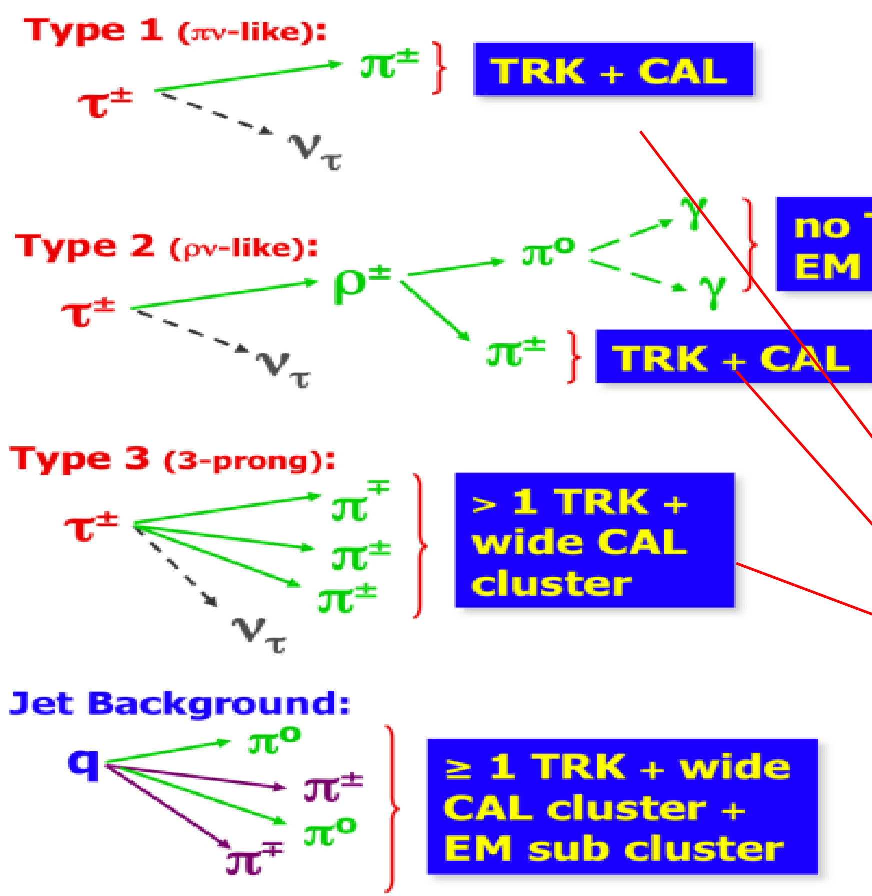
Tau categories:

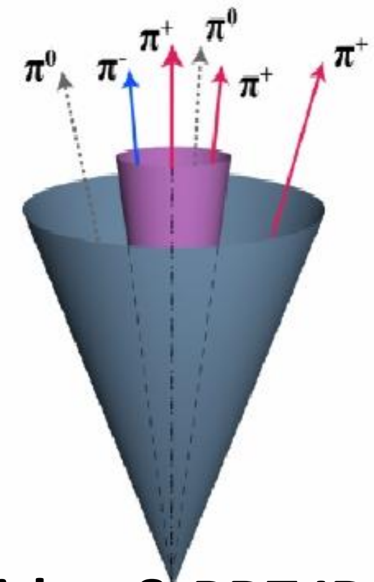
-1 prong ($\tau^{\pm} \rightarrow \nu_{\tau} + \pi^{\pm}$)

-1 prong + neutral pion ($\tau^{\pm} \rightarrow \nu_{\tau} + \pi^{\pm} + \pi^{0}$)

-3 prongs ($\tau^{\pm} \rightarrow \nu_{\tau} + 3\pi^{\pm} + (n)\pi^{0}$)

Hadronic tau identification





CDF: two cone algorithm & BDT ID

Tau categories:

-1 prong $(\tau^{\pm} \rightarrow \nu_{\tau} + \pi^{\pm})$

-3 prongs $(\tau^{\pm} \rightarrow \nu_{\tau} + 3\pi^{\pm} + (n)\pi^{0})$

Different energy subsamples trained separately

Strategies for the analysis 1

SIGNATURE SEARCH: similar approaches for CDF and DØ

looking for the leptonic+hadronic di-tau decay modes.

good compromise between:
- high hadronic decay B.R.
- good background rejection provided by e/μ identification

jets in the final state optimize sensitivity for $qqH \rightarrow qq\tau\tau$, $WH \rightarrow qq\tau\tau$ and $ZH \rightarrow qq\tau\tau$. $gg \rightarrow H$ events with jets from ISR are also included

CDF

One isolated lepton (e/μ) $p_T > 10$ GeV/c
One hadronic tau $p_{TVIS} > 15$ GeV/c
Opposite charges
$\geq$1calorimeter jet (DR=0.4 cone):
- $E_T > 20$ GeV
- EM fraction < 0.9
- pseudorapidity: $|\eta| < 2.5$

DØ

One isolated muon $p_T > 15$ GeV/c
One 1(3)-prong had. tau $p_{TVIS} > 15(20)$ GeV/c
Opposite charges
$\geq$2calorimeter jets (DR=0.5 cone):
- $E_T > 20$ GeV
- pseudorapidity: $|\eta| < 3.4$

Strategies for the analysis 2

BACKGROUND ESTIMATION

IRREDUCIBLE PHYSICS CONTRIBUTIONS

$Z\rightarrow\tau\tau$, top-antitop, dibosons : from MC

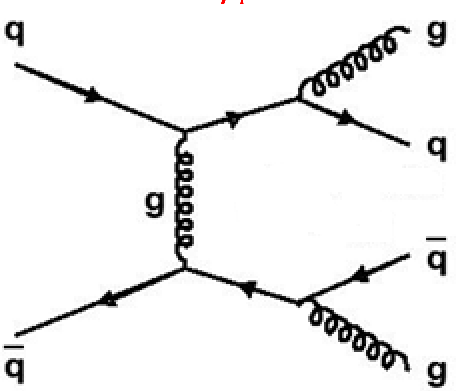
BACKGROUND FROM MISIDENTIFIED LEPTONS

W+jets, γ+ jet,multijet: based on MC and data driven techniques

THE CHALLENGE: evaluate jet$\rightarrow\tau$ fake rate. Extremely difficult.

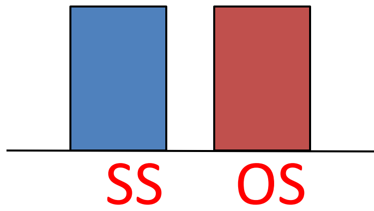
To estimate multijet bkg, both CDF and DØ use **same-sign (SS) data**:

$Q_\tau \times Q_{e/\mu}=1$



jet $\rightarrow$ e/μ
jet $\rightarrow\tau$

Little or no correlation is expected between charges

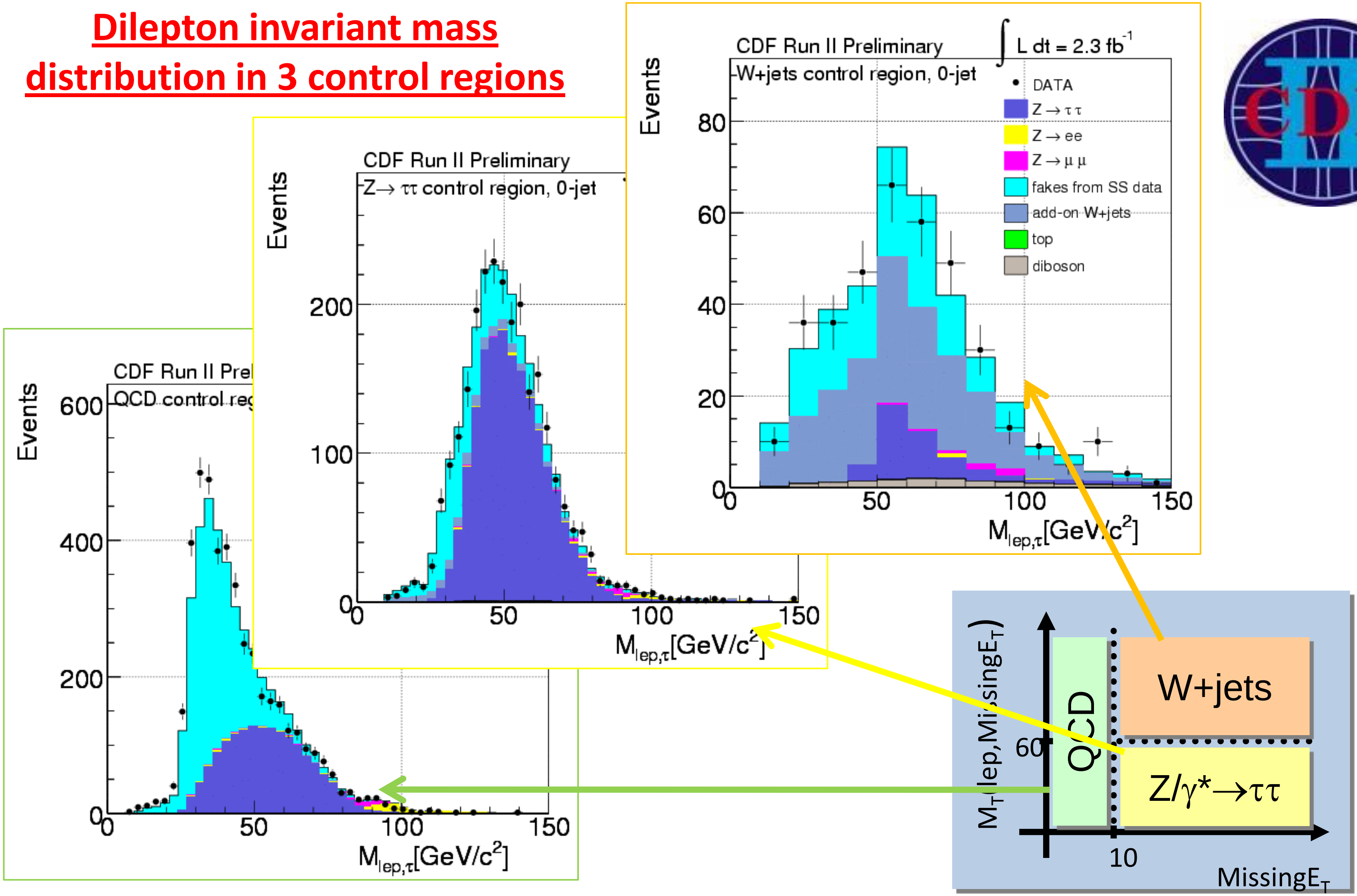


From isolation sidebands control regions:

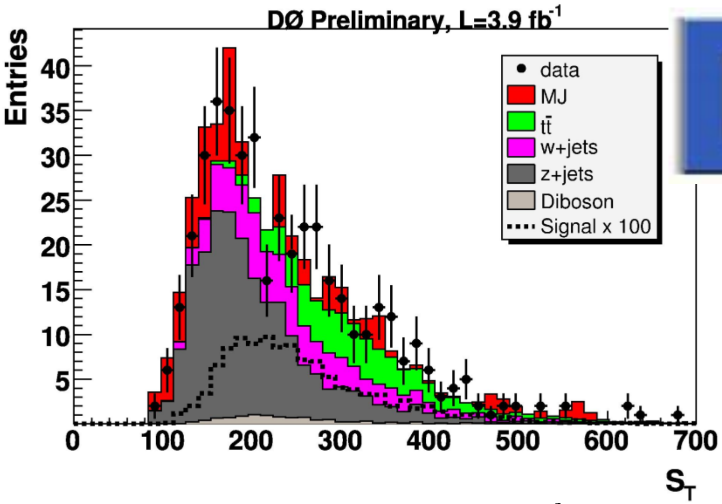
- Normalization factors for SS ~ 1
- Corrections for W+jets OS/SS asymmetries

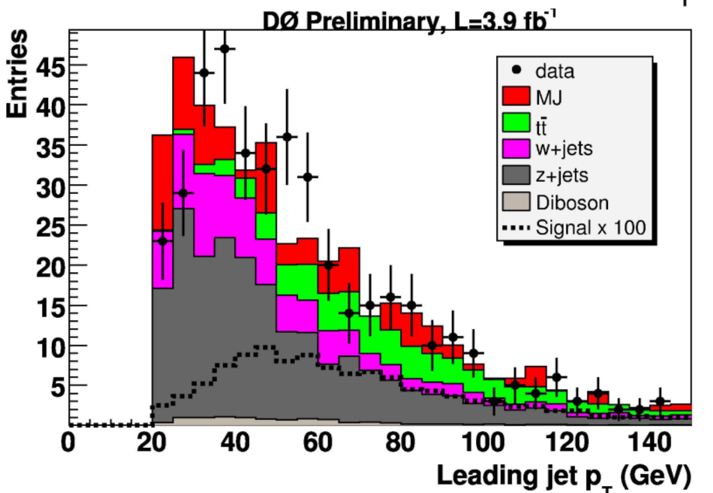
0-jet control region: background testing

Dilepton invariant mass distribution in 3 control regions



Signal channel: $\geq$ 2 jets





Data	ΣBknd	$t\bar{t}$	W+jets	Z+jets	DB	MJ
433	439.9	66.7	81.5	222.7	10.2	80.7

Process	ZH	HZ	HW	VBF	GGF
Event yield	0.11	0.23	0.72	0.12	0.15

Source	Uncertainty (%)
Luminosity	6.1
μ ID, track match, iso.	5.0
trigger	5.0
W/Z+light flavor XS	6.0
W/Z+heavy flavor XS	20.0
$t\bar{t}$, single top XS	10.0
diboson XS	7.0
Higgs boson XS	6.0
τ ID NN	8.9
Jet ID/reco eff.	3.0
Jet E resolution.	5.0
JES	7.5
jet p_T	10.0
pdfweight	Shape (currently 3.0)
MJ estimation	17

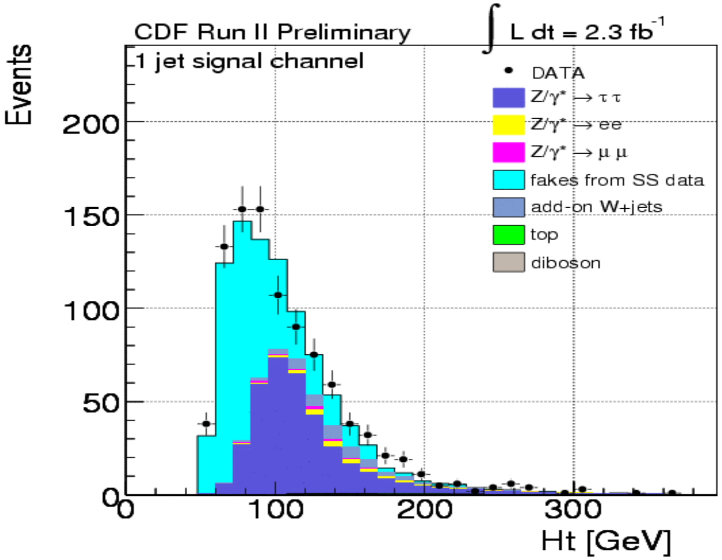
Main background Contributions:

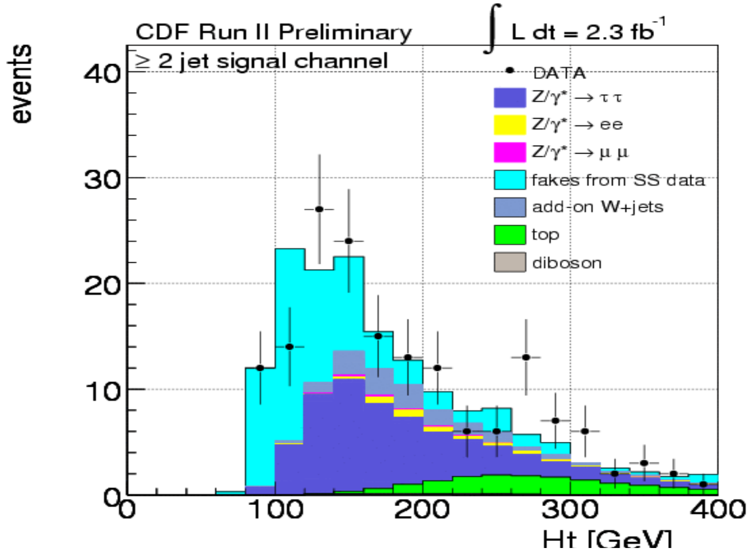
- $Z \rightarrow \tau\tau$
- jet$\rightarrow\tau$ fakes in QCD multijet and W+jets

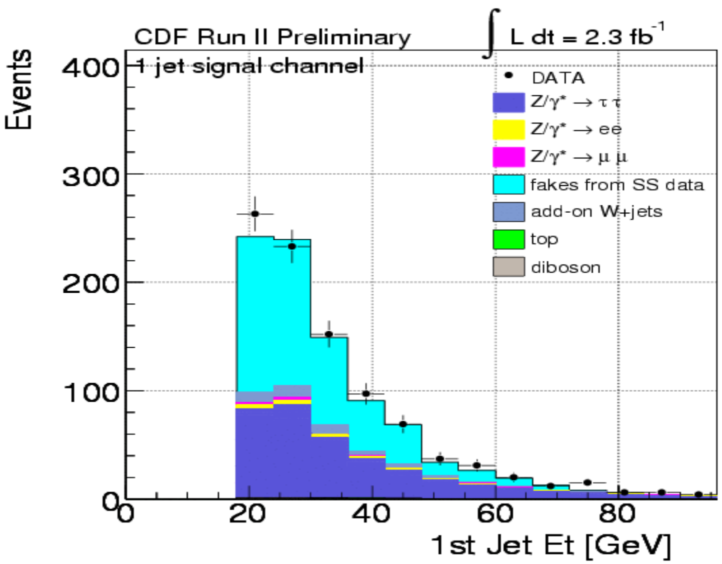
This search relies on a good jet multiplicity modeling.

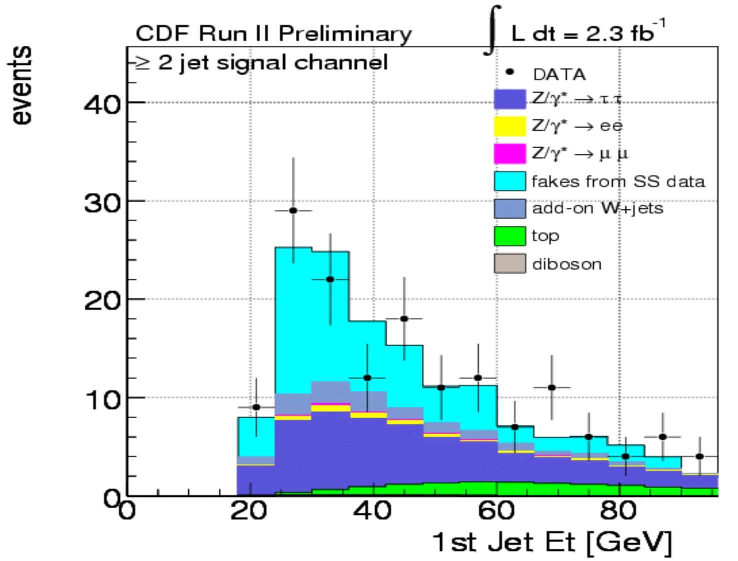
Thus, one of the main sources of systematics for MC-derived processes which has been considered is the uncertainty on the the **Jet Energy Scale (JES)**

Signal channels: 1 jet and ≥ 2 jets









Signal channels: 1 jet and ≥ 2 jets

	Data	Σ Bknd	$t\bar{t}$	add-on W+jets	$Z \to \tau\tau$	$Z \to ll$	DB	fakes from SS
1 jet	965	921.7	4.6	45.8	357.9	26.4	3.9	483.0
≥2 jets	166	159.4	16.3	14.1	59.3	4.8	0.9	64.0

	HZ	HW	VBF	GGF
1 jet	0.050	0.091	0.070	0.535
≥2 jets	0.099	0.150	0.099	0.129

Main background Contributions:

-jet→τ fakes in QCD multijet and W+jets

- Z→ττ

Source	Uncertainty (%) 1jet	Uncertainty (%) ≥2 jets
JES Drell-Yan	+6.2	+14.2
JES $t\bar{t}$	−7.7	+3.2
JES WW/WZ/ZZ	+7.1	+11.7
XS Drell-Yan	+2.2	+2.2
Acc.Drell-Yan	+2.3	+2.3
XS $t\bar{t}$	+10.0	+10.0
XS WW/WZ/ZZ	+6.0	+6.0
PDF bkgs	+1.0	+1.0
SS data	+10.0	+10.0
SCALE W+jets	+18.0	+30.0
tau ID:N_{obs}	+2.8	+2.8
tau ID:N_{SSdata}	−3.3	−3.3
tau ID:N_{Wjets}	−0.3	−0.3
tau ID:XS Drell-Yan	−2.1	−2.1
tau ID:Acc.Drell-Yan	−2.2	−2.2

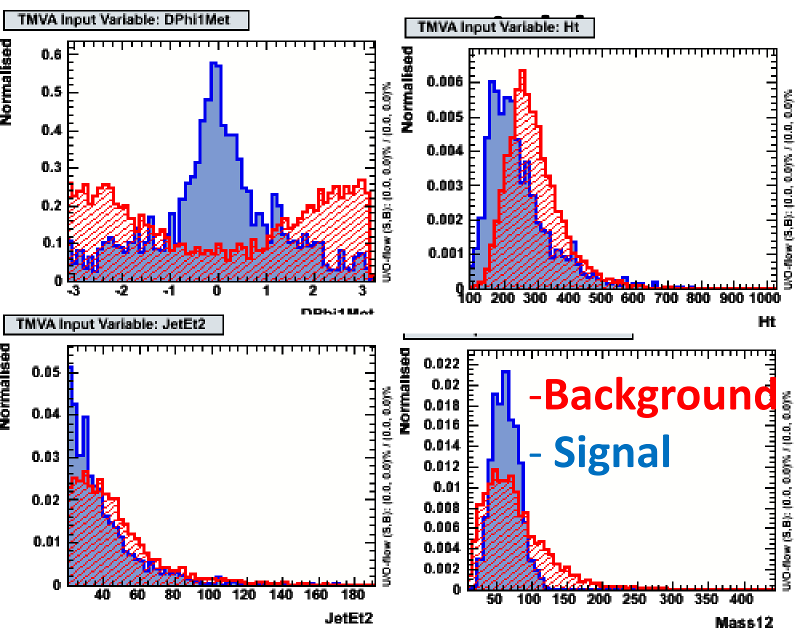
Jet Energy Scale (JES) uncertainty may affect <u>good jet multiplicity modeling</u>.

Systematic uncertainties related to the BDT-based Tau ID algorithm, evaluated in the 0-jet C.R.

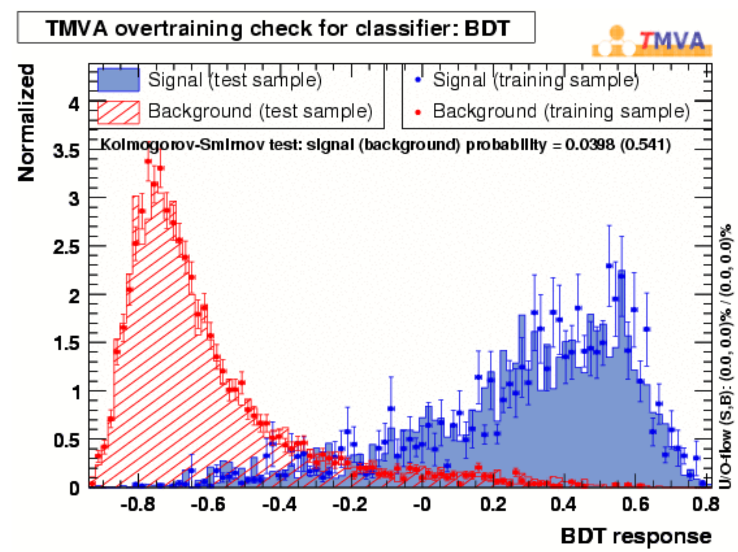
Signal vs. Background discrimination

- Good agreement in almost all kinematic distributions
- Expected signal is much smaller than background uncertainties
- S/B is small $\rightarrow$ counting experiment is not possible.
- Need to exploit all the event information to extract S from BKG

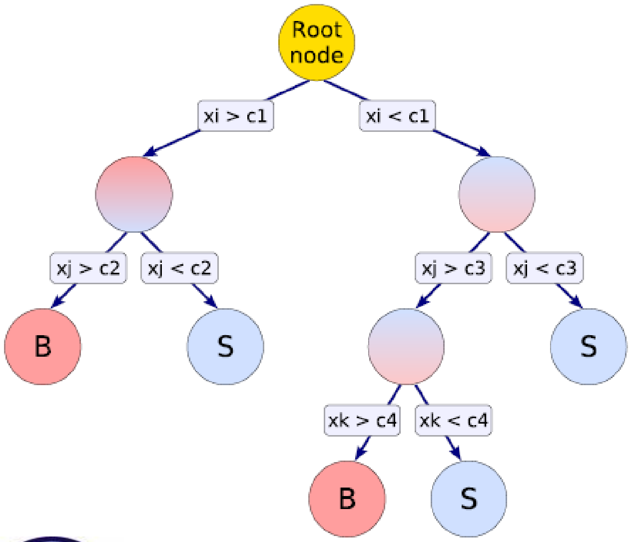
  A **multivariate technique** allows us to combine the **discriminating power** of different kinematical and topological distribution into



MULTIVARIATE ALGORITHM



Multivariate techniques

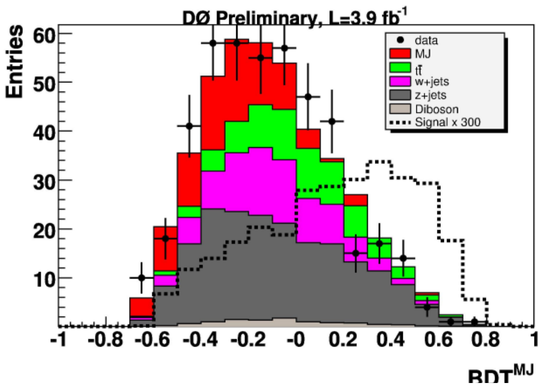


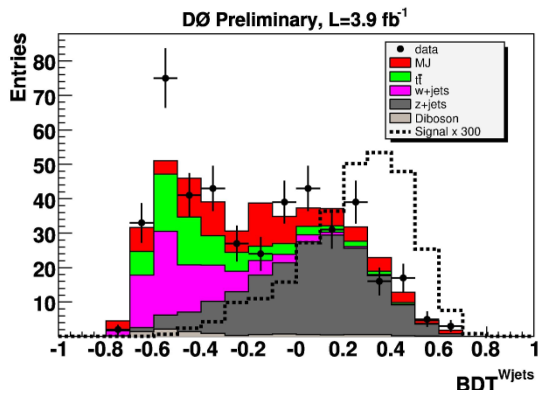
Both DØ and CDF employ a multivariate technique based on the **BOOSTED DECISION TREE** method

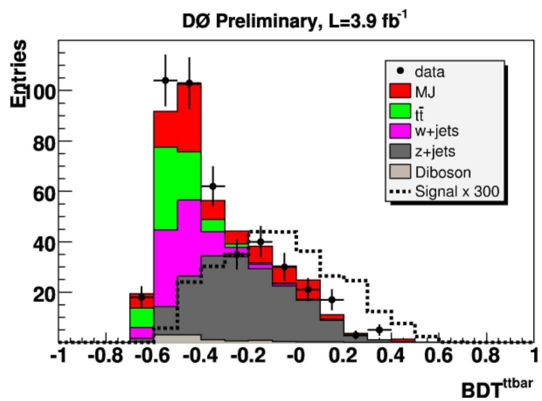
-train a mixture of signal processes vs different backgrounds: **top, Z+jets, multijet**
-combine outputs to maximize Signal vs Background rejection
- 1 jet and 2 jets separately

- train each signal process vs different backgrounds: **top, Z+jets, W+jets, multijet**
-combine outputs to maximize Signal vs Background rejection

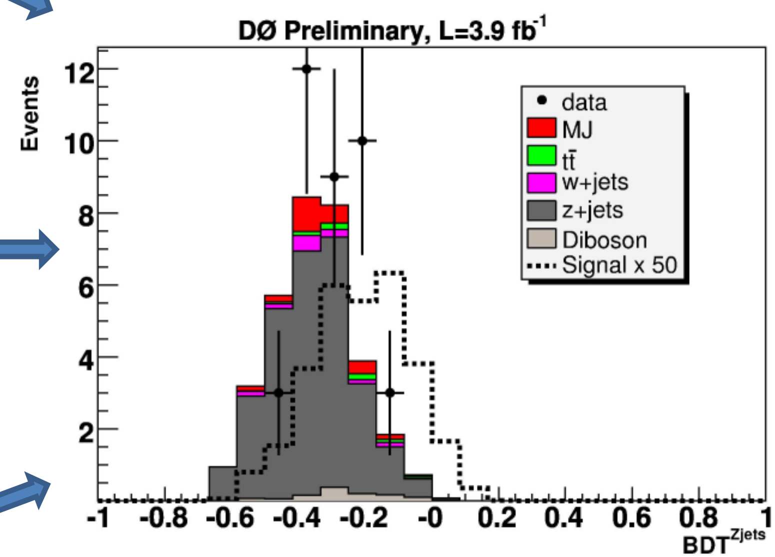
Multivariate discriminants







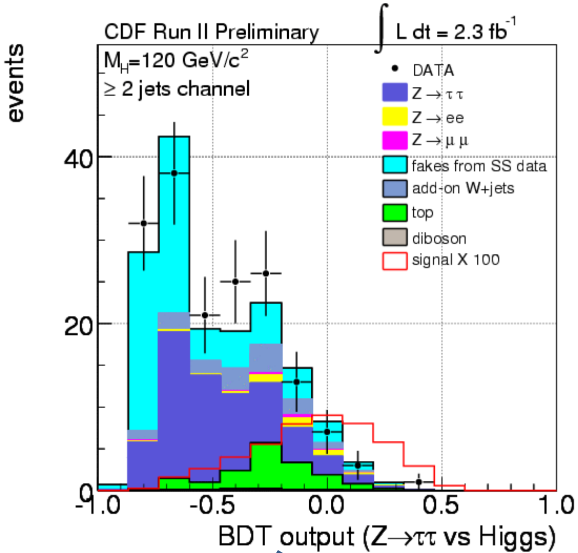
≥2 jets channel

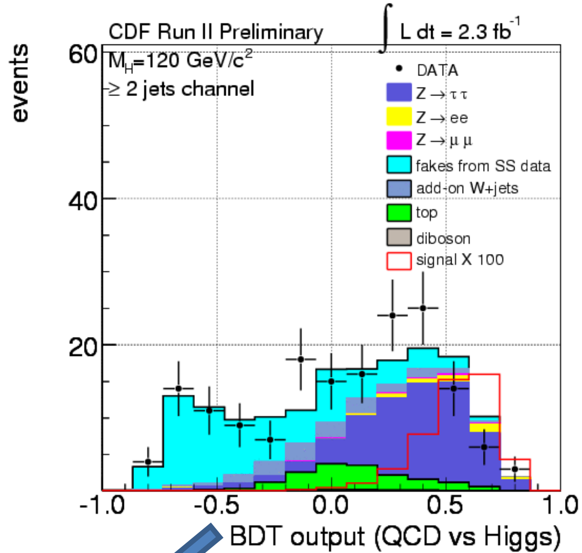


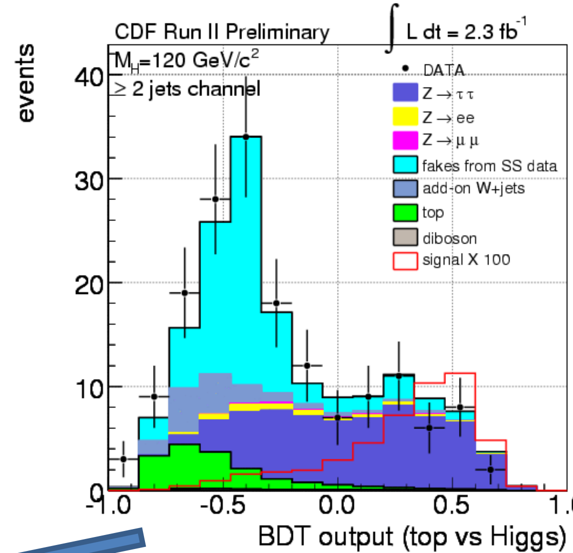
-Applying a lower cut on the BDT(S,MultiJet), BDT(S,W+JETS), BDT(S,tt)

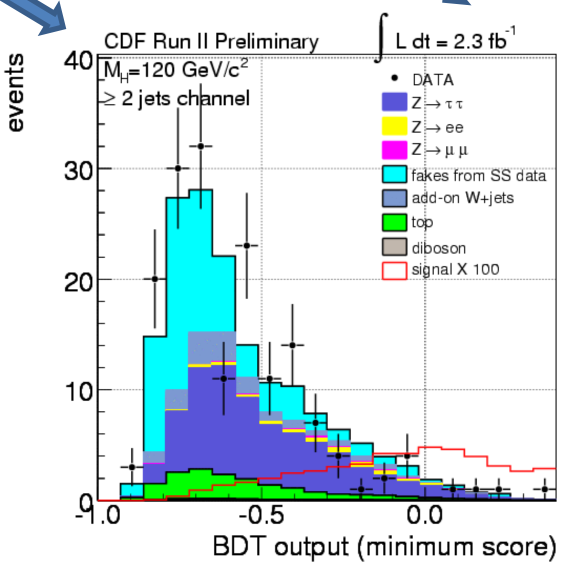
- Final discriminant: BDT(S,$Z\rightarrow\tau\tau$)

Multivariate discriminants





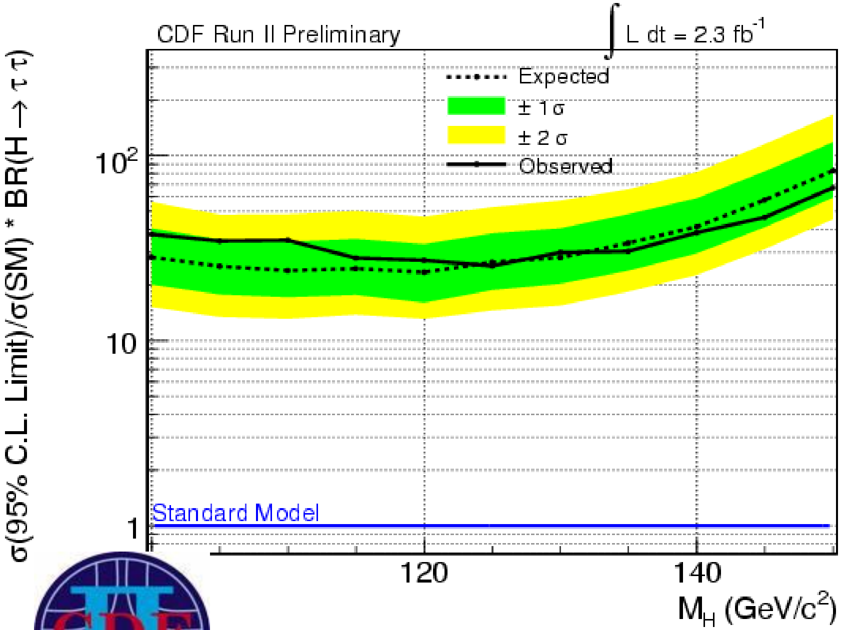




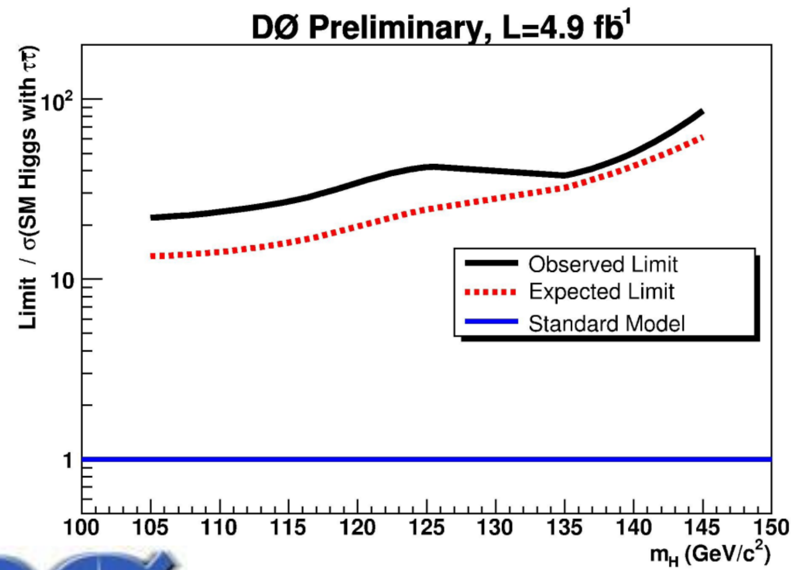
≥2 jets channel

-Final discriminant: taking the minimum output among BDT(S,$Z\rightarrow\tau\tau$), BDT(S,QCD), BDT(S,tt)

Results: 95% C.L. upper limit



2.3 fb^{-1}



4.9 fb^{-1}

Mass ranges explored:
100 – 150 GeV/c^2

CDF Expected limits: 23.4 – 82.6
CDF Observed limits: 25.3 – 70.0

Mass ranges explored:
105 – 145 GeV/c^2

DØ Expected limits: 13.4 – 61.4
DØ Observed limits: 21.9 – 86.0

Summary

• We presented the latest results of the SM Higgs searches at the Tevatron in the di-tau decay channel

•These analyses are aimed at completing the Higgs decay modes explored by CDF and DØ, with the purpose of increasing the experiment sensitivity in the low mass Higgs region

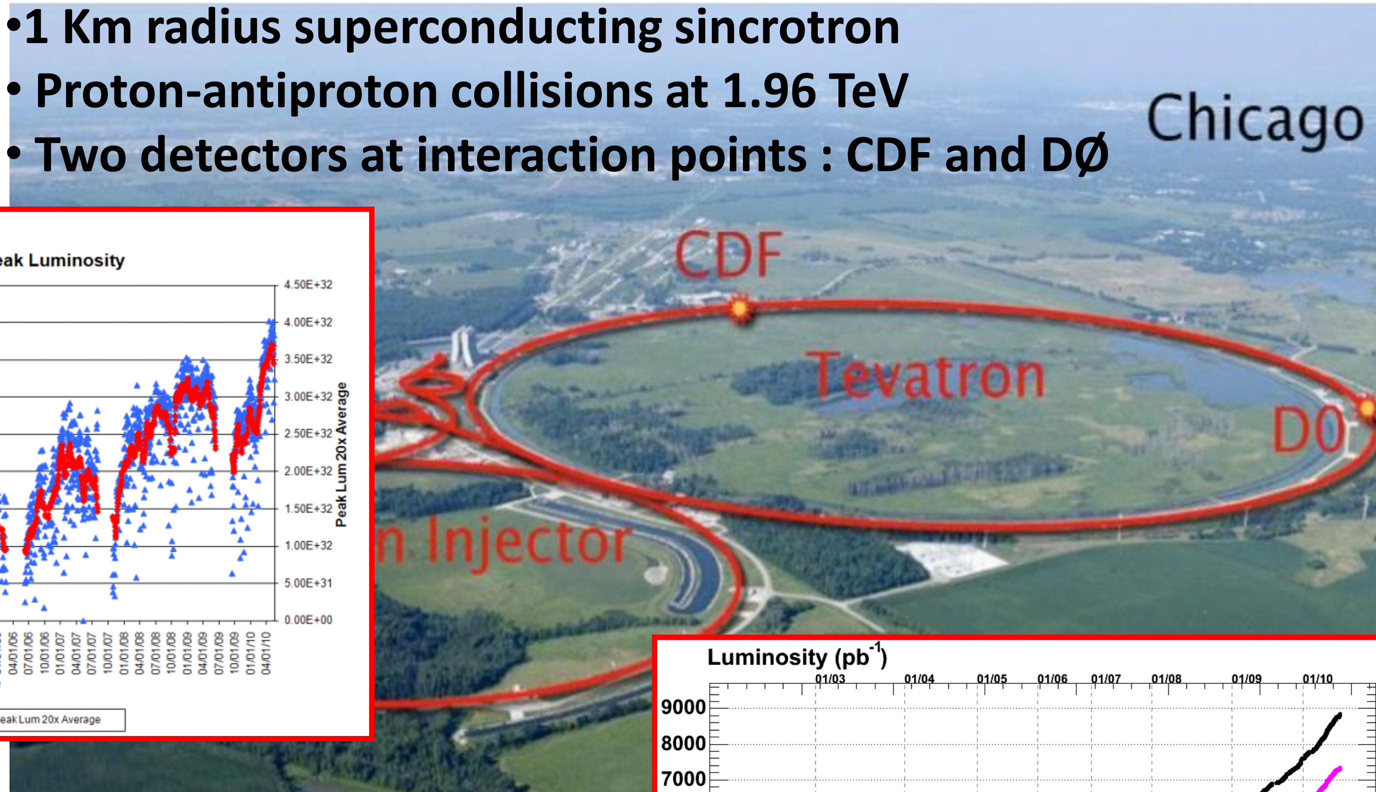
• CDF: 2.3 fb^{-1} expected(obs.) limit @ M_H = 115 GeV/c^2 24.5(27.9)
• DØ: 4.9 fb^{-1} expected(obs.) limit @ M_H = 115 GeV/c^2 15.9(27.0)

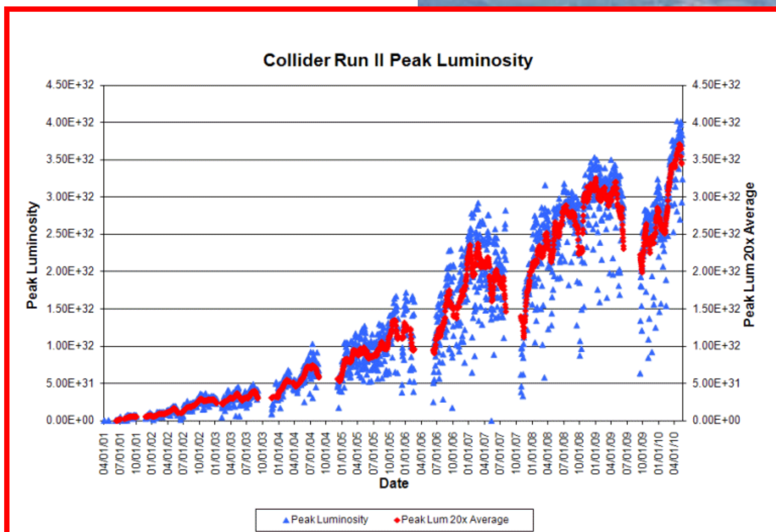
• Many improvements beyond luminosity scaling have been introduced since the previous stage of the analyses: new tau identification algorithms, increased acceptances, more sophisticated multivariate methods…
• Still working to add more data and get further improvements!

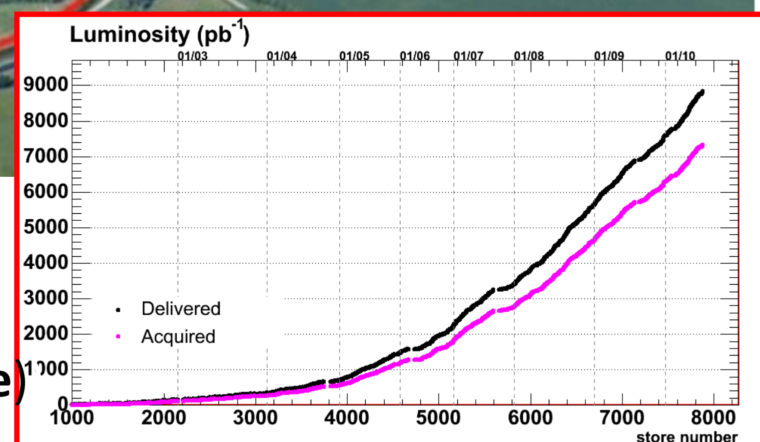
BACK-UP SLIDES

The Tevatron

- 1 Km radius superconducting sincrotron
- Proton-antiproton collisions at 1.96 TeV
- Two detectors at interaction points : CDF and DØ

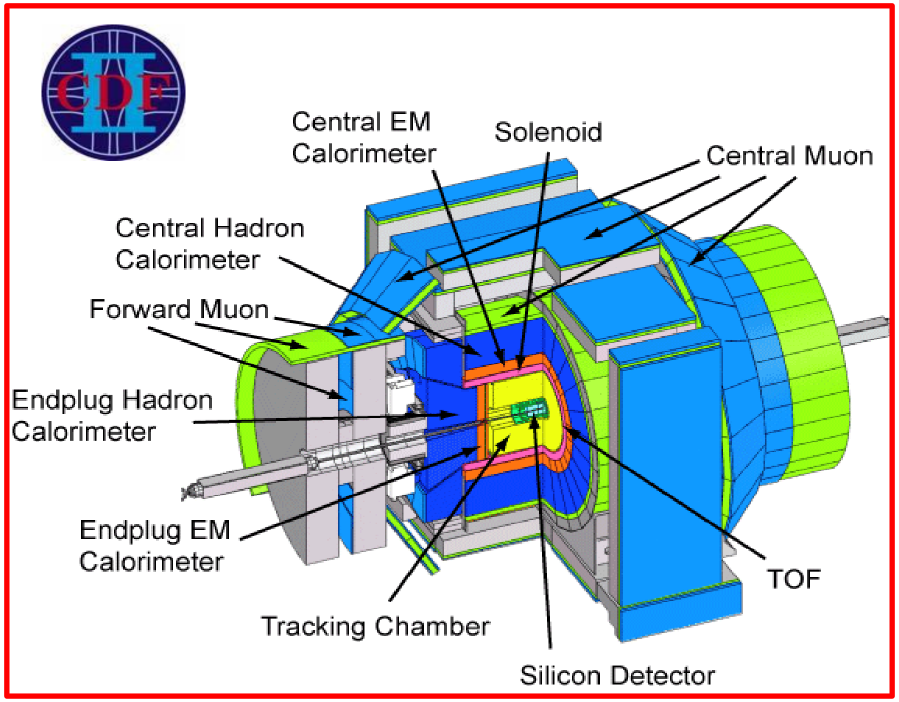






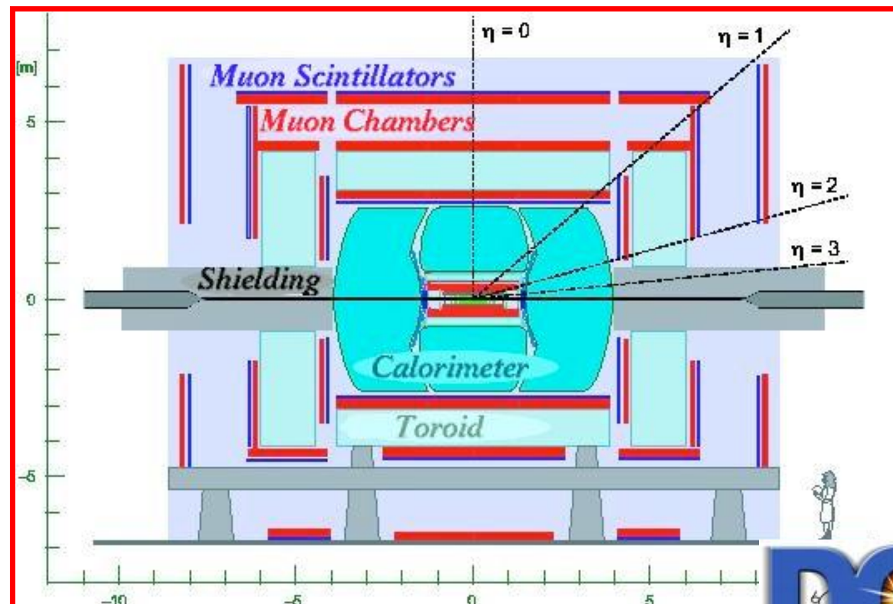
- peak luminosity 4 X 10^{32} cm^{-2}s^{-1};
- weakly integrated lum. ~60 pb^{-1};
-8.8 fb^{-1} delivered per experiment (7.4 fb^{-1} on tape)

CDF and DØ detectors

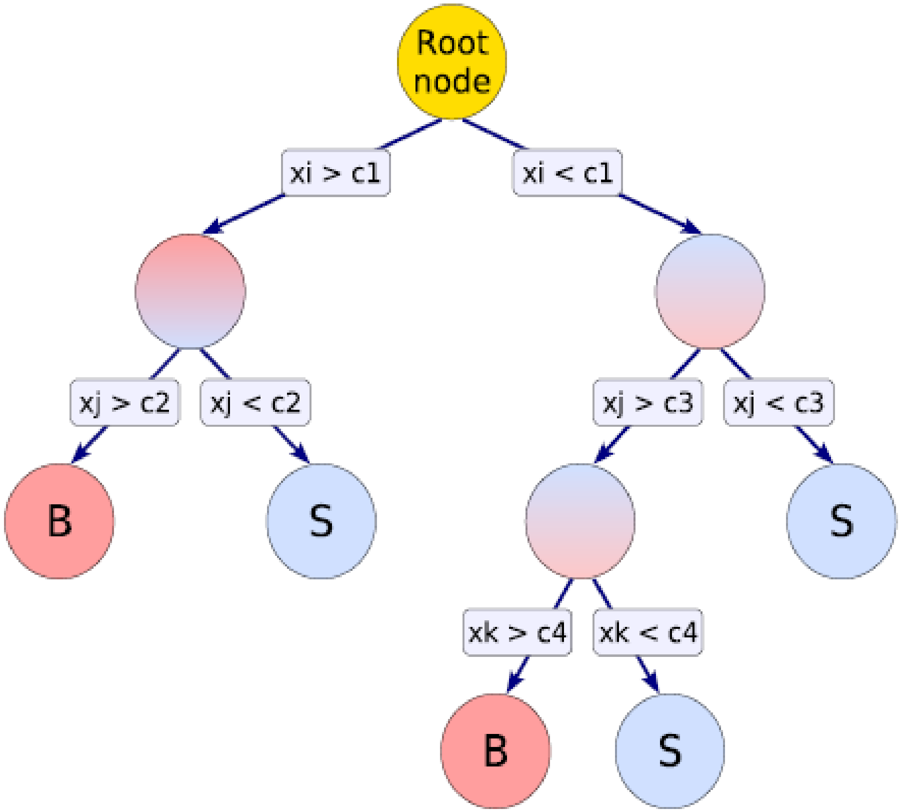


- Silicon Tracking |η|<2-2.5
- Drift cell Tracker 1.4 T, |η|<1.1
- Scintillator Cal. |η|<3.2
- Muons: |η|<1.5

- Silicon tracking |η|<3
- Fiber tracker 1.9 T, |η|<1.7
- LAr/DU calor. |η|<4
- Muons: |η|<2



The Boosted Decision Tree method



A DECISION TREE: a sequence of rooted binary splits

Ingredients : 1) a **<u>training sample</u>** for signal and background
2) a set of **<u>discriminating variables</u>**

At the end of a splitting, leaves are classified as signal-like (event score +1) or background-like (event score -1), accordingly to the purity.

BOOSTING: N trees are created. Events misclassified in the N-th tree, are given an <u>increased weight</u> in the (N+1)th tree.

An event final score is given by the weighted average of different tree outputs